DOCUMENT RESUME

ED 183 565 TM 010 007

AUTHOR Steinmetz, Andres
TITLE The Discrepancy Evaluation Model. II. The Application of the DEM to an Educational Program.
INSTITUTION National Council on Measurement in Education, East Lansing, Mich.
PUB DATE 76
NOTE 7p.; For related document, see TM 010 009
AVAILABLE FROM National Council on Measurement in Education, 1126 16th Street N.W., Washington, DC 20036 ($1.50)
JOURNAL CIT NCME Measurement in Education; v7 n2 Spr 1976

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Data Collection; Differences; Evaluation Criteria; *Evaluation Methods; Higher Education; *Input Output Analysis; Inservice Teacher Education; Models; Objectives; Performance; *Program Evaluation; Program Planning; Standards
IDENTIFIERS *Discrepancy Evaluation Model

ABSTRACT

The discrepancy evaluation model (DEM) specifies that evaluation consists of comparing performance with a standard, yielding discrepancy information. DEM is applied to programs in order to improve the program by making standards-performance-discrepancy cycles explicit and public. Action-oriented planning is involved in creating standards: a useful way to accomplish this is to analyze each of the components and subcomponents of the program, and then to subject the subcomponents to an input-output analysis. A "design evaluation" should also be undertaken to judge the adequacy of program intentions. Programs plans may serve as the basis for the DEM's standards; process evaluation determines whether and how well planned activities are carried out; and outcome evaluation determines the extent to which planned outcomes (both subobjectives and terminal or final objectives) are achieved. Invariably, priorities must be set from the initial list of planning objectives, and the evaluator will not study planning objectives unless a standard exists for evaluating them. (DEM is briefly applied to an evaluation of a teacher inservice education program as an example, and a chart is provided to summarize data collection planning). (GDS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NCME

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED183565

TM010 002

measurement in education

A NEWS OF SPECIAL INTEREST TO THE EDUCATIONAL MEASUREMENT COMMUNITY

The Discrepancy Evaluation Model

Andrés Steinmetz
President, New Measures, Inc.

II. THE APPLICATION OF THE DEM TO AN EDUCATIONAL PROGRAM

Review of the Model

In section I (Vol. 7, #1) we said that evaluation always consists of comparing **Performance (P)** with a **Standard (S)**. This comparison yields **Discrepancy (D)** information which can be used as a basis for making a judgment of value or worth about the object being evaluated.

To evaluate something we must have a pertinent **S** available. Obtaining such an **S** is usually not easy and in most circumstances has to be created — a job done by the client, assisted by the evaluator, to clarify and make conscious the **S** that should govern the activity or object being evaluated. Usually the extent to which **S** can be made explicit and observable is a matter of degree — important dimensions of it remaining implicit.

ABOUT THIS REPORT

As indicated in volume 7, Number 1, this is the second in a two-part series on the Discrepancy Evaluation Model (DEM). In this section, Dr. Andrés Steinmetz reviews briefly the assumptions of the model and then proceeds to illustrate the application of the DEM to an educational program. Specific procedural suggestions are included as part of this presentation.

Dr. Steinmetz is currently president of New Measures, Inc., a private consulting firm in the areas of organizational development and evaluation, Charlottesville, VA, 22903.

PSR

The evaluator collects data for which an explicit **S** is available. But in order to do so, the evaluator and client must first agree on both the specific **P** information to be collected and the source of that information. This may involve the client and evaluator working together to continually clarify **S**.

Basic Steps in the Application of the DEM to a Program.

The first thing to do is to understand that the purpose of the evaluation is to improve the program by making **SPD** cycles explicit and public insofar as is possible. This includes agreeing to the role distinctions between client and evaluator discussed in Section I (Vol. 7, #1) and clearing the way for the first task which is to create **S**.

Creating **S** is action-oriented planning. The client must turn both to existing knowledge and to his/her own experiences, values, and purposes in order to construct **S**. He/She must seek to involve others on his/her staff, those affected by the program, or those for whom the program is designed, in order to end up with an adequate and realistic **S**. Creating **S** is thus very much an exercise in applied goal and value clarification and may be thought of as creating a concrete model of a program.

A useful way to proceed in order to create **S** is to do a component analysis: to break the program into its major activities, functions, or components. Each component in turn can then be broken into its subcomponents and so on until a level of detail is reached suitable to the needs of program management. For example, suppose the program we are concerned about is a teacher in-service program. We can represent it as shown in Figure 1. However, upon thinking through the basic organization, we might decide that the program really consists of three major components: selection, curriculum development, and instruction. These can then be

represented as subcomponents of the teacher in-service program as shown in Figure 2.

Figure 1
Level I Analysis

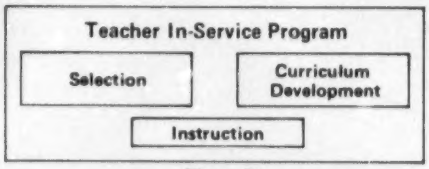


Figure 2
Level II Component Analysis

Each time we break a component into subcomponents, we reach a new "level" of detail. Thus, we call Figure 1 Level I analysis and Figure 2 Level II analysis. Figure 3 shows a Level III analysis for the Selection component. While Selection alone has been chosen for the sake of brevity, the other components would be similarly broken down when constructing a program design.

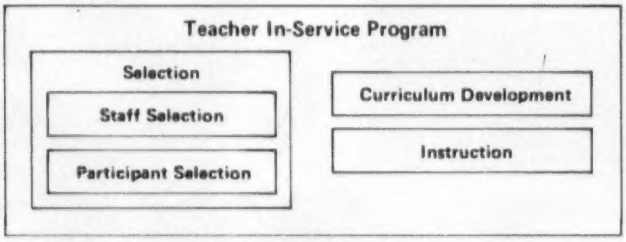
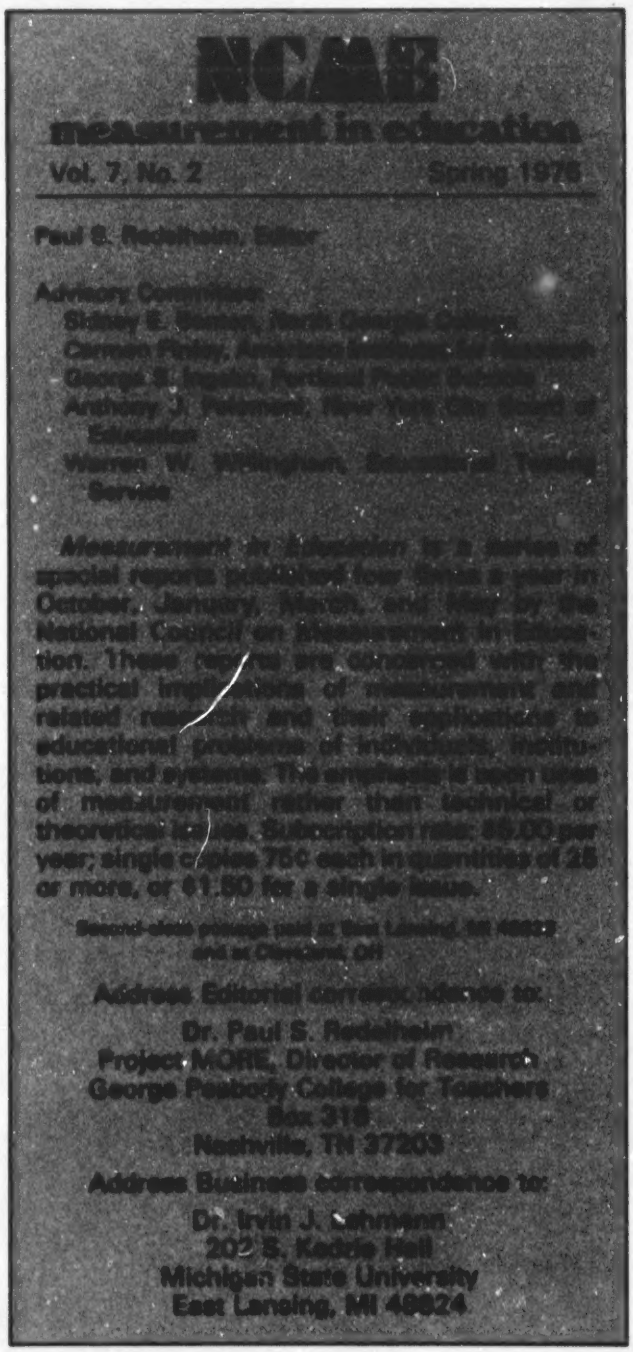


Figure 3
Level III Component Analysis of the
Teacher In-Service Program

In order to write a practical description of each component and subcomponent, we do an input-process-output analysis for each. This means that we assume program activity is not random, that it is goal directed, and that each activity has one or more objectives. These objectives, which may be conditions, behaviors, tangible products, or any purpose an activity is trying to realize, are *outputs*. The things we do to bring about the outputs are *processes*. Processes indicate what will be done, who will do it, how, when, and where. They describe how resources will be combined or transformed to produce outputs. The resources themselves, the personnel, facilities, materials, prerequisites, etc., that are needed to support the processes, are *inputs*.

Let's assign the entire program the numeral 1.0. We could then call the selection component 1.1, curriculum development 1.2 and instruction 1.3. Similarly, "staff selection" can be 1.1.1. and "participant selection" 1.1.2. (This numbering system offers a convenient way to refer to components and follows the usual outlining form.) An abbreviated input-process-output description for each might look as shown in Table 1.

In the same way, a program design or **S** would be developed for each of the other subcomponents, breaking these down further into their components (thus, going to Level IV detail) if that is useful. Notice that the output of 1.1.1, "two teacher trainers" is referenced "to 1.1.2, 1.2, and 1.3." This indicates that the output of 1.1.1 is used as input to 1.1.2, 1.2, and 1.3. In other words, the teacher trainers will be inputs to the participant selection, curriculum development, and in-service instruction activities. The input-process-output description for 1.1.2 shows this relationship between 1.1.1 and 1.1.2 with the entry "teacher trainers (from 1.1.1)" in the input column. We would thus expect to find the


NCME
measurement in education
Vol. 7, No. 2 Spring 1978

Paul S. Redelheim, Editor

Advisory Committee:
Sidney E. [illegible], North Carolina State University
Carmen Finley, Assistant Superintendent for Research, Portland Public Schools
George S. [illegible], Portland Public Schools
Anthony J. Nitko, University of Pittsburgh, Graduate School of Education
Warren W. Willingham, Educational Testing Service

Measurement in Education is a series of special reports published four times a year in October, January, March, and May by the National Council on Measurement in Education. These reports are concerned with the practical implications of measurement and related research and their application to educational problems of individuals, institutions, and systems. The emphasis is upon issues of measurement rather than technical or theoretical issues. Subscription rate: $5.00 per year; single copies 75¢ each in quantities of 25 or more, or $1.50 for a single issue.

Second-class postage paid at East Lansing, MI 48823 and at Cleveland, OH

Address Editorial correspondence to:
Dr. Paul S. Redelheim
Project MORE, Director of Research
George Peabody College for Teachers
Box 318
Nashville, TN 37203

Address Business correspondence to:
Dr. Irvin J. Lehmann
202 S. Kedzie Hall
Michigan State University
East Lansing, MI 48824

INPUT	PROCESS	OUTPUT
–Program director –3 principals –3 building coordinators –teacher needs assessment data –$30,000 personnel budget –Personnel Office procedures	**1.1.1 STAFF SELECTION.** The program director meets with the principals and building coordinators of each of the three schools in order to: –write job descriptions for two full-time teacher trainers based on needs assessment data; –coordinate with the school district personnel office and their recruitment procedures; –plan the applicant screening and selection procedures; –carry out the selection procedure.	–two teacher trainers (to 1.1.2, 1.2, 1.3).
–Program director –teacher trainers (from 1.1.1) –teacher needs assessment data –3 building coordinators	**1.1.2 PARTICIPANT SELECTION.** (Describes how the participants for the in-service program will be selected.)	–10 teachers from each school (to 1.3).

Table 1
Input-Process-Output Description for Two Components

contribution of the teacher trainers mentioned in the description of the process for 1.1.2, Participant Selection. The design also shows that the ten teachers from each school are inputs to component 1.3, In-Service Instruction (that makes sense, of course, since they will be the beneficiaries of the program).

These input to output relationships are shown in the network in Figure 4 which is the same as Figure 6 except that the component numbers and some arrows between components have been added. An arrow between components means that one component produces at least one output which is an input to the other component. Thus, the fact that teacher trainers selected in 1.1.1 are an input to 1.1.2, as already noted, is shown with an arrow connecting 1.1.1 and 1.1.2. The other (input-output) relationships among components are also shown with arrows. This additional information to the component analysis of Figure 3 leads us to call Figure 4 a "network."

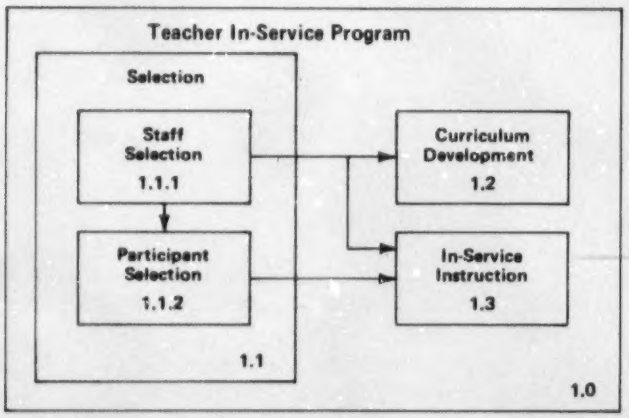


Figure 4
Level III Program Network

This kind of component and relationship analysis eventually produces a detailed program design. It will consist of input-process-output narratives for each component along with a network showing all components (and subcomponents) and the major relationships among them. This design then acts as **S**; specifying what *should* be, the intent of the program. As we will see in a moment, **P** data can then be collected on any aspect of the program to determine whether what should be occurring or resulting **(S)** is *actually* taking place **(P)**.

It should be evident that this planning procedure stands to surface differences in values, approaches, procedures, philosophical orientations, etc. Properly facilitated by the DEM evaluator, planning promotes resolution or negotiation of these differences and agreement on the **S** that will govern formal evaluation. As these plans are implemented and periodically reviewed in light of **P** information, they may be changed or amended — that is, the **S** may change as conditions change and as continuous assessment of results occurs. Thus, programmatic decisions and actions utilize feedback. The evaluator's job includes gathering that feedback and putting it at the disposal of the evaluator's client(s).

The input-process-output description suggests different kinds of information that may be deliberately or formally gathered in order to assist feedback-guided action. These are summarized in Table 2.

"Design evaluation" in Table 2 refers to judging the adequacy of program intentions. The object being evaluated here is the program plan. Any program is bound to have a basis in social and moral philosophy, as well as empirical research, and the critique of these bases may be referred to as the problem of construct validity. But the program plans may also be examined for their comprehensiveness, appropriateness to the situation, relationship to

MANAGEMENT ACTION	KINDS OF EVALUATION	STANDARDS RELATIVE TO:
Program Planning	Design Evaluation	—design adequacy; construct validity, comprehensiveness, internal or logical consistency, relationship to need, appropriateness.
Program Implementation	Input Evaluation	—availability of resources; extent to which prerequisites and preconditions are met, extent to which program resources ($, people, materials, etc.) are available and deployed as planned.
	Process Evaluation	—extent to which activities are carried out as planned; existence or frequency, intensity, manner, and other qualities and characteristics of the activities.
	Outcome Evaluation	—interim outcomes; outcomes which are prerequisite to other outcomes; terminal outcomes.

Table 2
Kinds of Evaluation Useful to Program Improvement

known interests and needs, and so on. An analysis may also be made to see whether resources such as the kinds and qualifications of personnel and materials seem adequate to support the activities which the program intends to undertake. Similarly, one may critique the logical relationship between program objectives and the activities designed to bring them about. Design evaluation, then, refers to the construct and logical or operational validity of a set of intentions. The standards involved in this sort of evaluation are often not entirely explicit in advance and are made explicit incrementally. The method used is that of logical agrument and the evaluation itself is readily understood in terms of the **S, P, D** concepts.

Program plans may themselves serve as **S** for other evaluation undertaken during the life cycle of the program. Program plans specify and direct program implementation, and as such, may serve as **S** to input, process, and outcome evaluation. For example, any program utilizes certain kinds and amounts of resources over time. Program plans which specify the number and kind of resources needed at different points in time relative to different activities and purposes can act as an **S** governing the installation of the program. Thus, **P** information may be gathered concerning the extent to which the resources planned for are indeed available, and are in fact deployed as required (input evaluation). A program whose design has been judged adequate may nevertheless falter if it does not have the proper resources available when they are needed. Input evaluation is aimed at helping management make sure these resources are available when necessary.

Process evaluation involves determining whether planned activities are carried out in the manner called for by the program plans and whether they are of the quality expected. Again, the **S** here is the program plan which specifies and describes the program processes to be set into motion. Because of the complex interaction between **S** and action (which was known first?), thorough process evaluation overlaps with action research.

Outcome evaluation refers to determining the extent to which planned outcomes are achieved. It is useful to distinguish at least two classes of outcomes. Enabling or interim outcomes refer to milestones or subobjectives essential to the execution of the program from month to month. In contrast, terminal objectives refer to the major purposes or aims of the program.

It should be clear that properly specifying inputs, processes, and outputs for each component and subcomponent and specifying the relationships among all subcomponents, amounts to making available the **S** essential for input, process, and outcome evaluation. This makes it possible to conduct evaluation on a continuous basis throughout the life of the project because **P** data can be gathered relative to a larger class of program characteristics than just terminal objectives. Finally, the **D** information produced in the course of evaluation may be used to support two broad categories of management action. It may be used, on the one hand, to exert greater control over program operations in order to insure that **P** meets **S**. On the other hand, management may decide that the **S** originally set is inappropriate or unrealistic and may thus change the **S** involved.

Now it is, of course, impossible to formally collect empirical information on *all* inputs, processes, and outputs. Thus, program management is faced with setting some priorities. Management must identify the **P** information that would be most useful to it given its limited resources and its internal and external needs. There will be **P** information useful primarily to program management in the day-to-day operation of the program and there will be information which has to be provided to individuals and other organizations in the environment that serves to justify the program. Thus, management must set priorities around its needs for proper internal management and its need to remain accountable to the external environment. The decisions involved

REPORTS AVAILABLE

Back issues of *Measurement in Education* are available at 50¢ each in quantities of 25 or more for a single issue.

Vol. 1, No. 1	*Helping Teachers Use Tests* by Robert L. Thorndike
No. 2	*Interpreting Achievement Profiles — Uses and Warnings* by Eric F. Gardner
No. 3	*Mastery Learning and Mastery Testing* by Samuel T. Mayo
No. 4	*On Reporting Test Results to Community Groups* by Alden W. Badal & Edwin P. Larsen
Vol. 2, No. 1	*National Assessment Says* by Frank B. Womer
No. 2	*The PLAN System for Individualizing Education* by John C. Flanagan
No. 3	*Measurement Aspects of Performance Contracting* by Richard E. Schutz
No. 4	*The History of Grading Practices* by Louise Witmer Cureton
Vol. 3, No. 1	*Using Your Achievement Test Score Reports* by Edwin Gary Joselyn & Jack C. Merwin
No. 2	*An Item Analysis Service for Teachers* by Willard G. Warrington
No. 3	*On the Reliability of Ratings of Essay Examinations* by William E. Coffman
No. 4	*Criterion-Referenced Testing in the Classroom* by Peter W. Airasian and George F. Madaus
Vo. 4, No. 1	*Goals and Objectives in Planning and Evaluation: A Second Generation* by Victor W. Doherty and Walter E. Hathaway
No. 2	*Career Maturity* by John O. Crites
No. 3	*Assessing Educational Achievement in the Affective Domain* by Ralph W. Tyler
No. 4	*The National Test-Equating Study in Reading (The Anchor Test Study)* by Richard M. Jaeger
Vol. 5, No. 1	*The Tangled Web* by Fred F. Harcleroad
No. 2	*A Moratorium? What Kind?* by William E. Coffman
No. 3	*Evaluators, Educators, and the Publics: A Detente?* by William A. Mehrens
No. 4	*Shall We Get Rid of Grades?* by Robert L. Ebel
Vol. 6, No. 1	*On Evaluating a Project: Some Practical Suggestions* by John W. Wick
No. 2	*Measuring Reading Achievement: A Case for Criterion — Referenced Testing and Accountability* by S. Jay Samuels and Glenace E. Edwall
No. 3	*Security in a Citywide Testing Program* by Anthony J. Polemeni
No. 4	*Dear Mama: Why don't they love me anymore?* by Dr. Thomas J. Fitzgibbon
Vol. 7, No. 1	*The Discrepancy Evaluation Model (Part I)* by Andrés Steinmetz

are made by management, not the DEM evaluator, although the latter again facilitates the deliberations involved. Having a complete program design available literally points out trouble spots and helps in making the trade-offs involved.

The collection of **P** information is guided by what the DEM evaluator calls "evaluation questions." Such questions ask whether what *should* be actually *is*; whether inputs are available *as specified*; whether processes are carried out *as planned*; and whether outcomes are being achieved *as intended.** In other words, evaluation questions direct attention to the **P** information needed in order to determine whether the applicable **S** has been met. Examples might be: Are there 10 participants from each school and do they meet selection criteria? (outcome evaluation; the **S** required that there be 10 from each school meeting certain criteria); Is the needs assessment data available? (input evaluation; the **S** required that needs assessment data be available to the people planning the selection of staff); or, Is the personnel committee meeting as planned? (process evaluation; the **S** specifies who should meet to design and carry out the staff selection). There might also be evaluation questions about the operation of the other components, similarly aimed at ensuring effective program operation. And, no doubt, there would be evaluation questions aimed at determining whether the terminal outcomes have been realized.

It is important to notice the very narrow definition given to evaluation questions. Evaluation questions assume the existence of an **S**. This is because evaluation is defined as the comparison between what *is* and what *might* be and is impossible unless the **S** is specified. The DEM evaluator will not let himself/herself get involved in collecting **P** information to answer questions for which no **S** exists. But, as already discussed, he/she will work with the client to articulate the **S** and to define the action that needs to be taken in order to make a pertinent **S** available.

The connection between the program design, the *intent* or expectations of the program, in short, the program **S**, and the program as it *actually is,* is provided by the evaluation questions. Dozens of these may be asked and many can be answered through informal interviews, meetings, or planning sessions. Some will be pursued more formally depending upon the interests and needs of management and the problems and cost involved in collecting the information. The major steps essential to the collection of **P** information are summarized in Tables 3 and 4. Thus, a DEM evaluation will consist of an **S** (a detailed program design showing a network and input-process-output descriptions for all components and subcomponents) and a data collection plan filled out for each evaluation question asked. In this way an internal feedback cycle can be set up so that the program is managed as much as possible on the basis of **D** information generated by comparing **S** and **P**, that is, on the basis of systematic evaluation.

---

*This does not necessarily mean that the DEM evaluator will not be open to unexpected events. How to handle this problem is negotiated between client and evaluator; the important thing, again, is for the evaluator to stay away from setting *S* and deciding on his/her own what information to collect.

(a) Evaluation Questions	(b) Design Referent	(c) Standard	(d) Data Collection Strategy	(e) Instrument	(f) Data Information Needed
(The specific question about the program and its operation that management wants answered. These questions give feedback important to determining whether Standards should be changed or whether greater control over Performance should be exerted. The Standards applicable to each evaluation question must be known.)	(Names the component, its number, and whether the evaluation question is dealing with an input, process, or output. This refers you to the design where the Standards governing the evaluation question are described.)	(Summarizes the Standards pertinent to the evaluation question. This information is extracted from the program design.)	(Lists sources of information and techniques for obtaining the information. Represents the thinking that must lead to the "instrument" in the next column and to Table 4.)	(A general term standing for the method or procedure to be used in gathering data; i.e., test, questionnaire, interview, observation, etc.)	(Dates are crucial particularly when information is supposed to support program planning and operation.)

Table 3
Summary of Data Collection Plan

(a) Instrument	(b) Evaluation Question Addressed	(c) Instrument Administration Schedule	(d) Data Collector	(e) Respondents	(f) Sample	(g) Data Analysis Procedure & Analyst	(h) Audience	(i) Report Preparation and Date
(From column (e) in Table 3.)	(Shows the number of each evaluation question for which the instruments gather data. Often any one instrument is designed to answer more than one question.)	(Shows data collection dates and times.)	(Shows who is responsible for administering the instrument or collecting the data.)	(Shows the direct source of data: people, files, etc.)	(Usually limited resources demand that the respondents be sampled.)	(Who will analyze data and how the analysis will be done must be thought through in advance.)	(All persons or places expecting the information gathered are listed.)	(Names the persons with primary report preparation responsibility and the date by which reports must be ready.)

Table 4
Summary of Data Collection Plan (continued)

III. CONCLUSION

The Discrepancy Evaluation Model offers a pragmatic, systematic approach to a wide variety of evaluation needs. From the daily activities of an individual teacher to educational program evaluation, the DEM can be utilized to structure the gathering of information essential for well-informed decision-making. A major feature of the DEM is its emphasis on self-evaluation and systematic program improvement.

NCME NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

202 S. Kedzie Hall
Michigan State University
East Lansing, Michigan 48824

Second class postage paid
at Cleveland, Ohio

A series of special reports of the National Council on Measurement in Education